

Planejamento da Análise de Dados

1) Perguntas de negócio para direcionar a análise:

Nesta etapa, é necessário pensar nas possíveis perguntas de negócio que envolvam a análise que busca responder a pergunta principal. Além disso, é importante já identificar os campos (da tabela) que serão usados para responder às perguntas.

Pergunta principal:

- **Onde é o melhor lugar para se investir em casas de temporada em Boston?**

Identificar a melhores épocas:

- *Quais são as épocas mais movimentadas? (pesquisar)*
- *Quais são os bairros de Boston que aparecem no dataset? (neighborhoods) (colocar os ícones no mapa para mostrar onde estão localizados)*

Identificar as casas a serem analisadas:

- *Quais são as casas com mais avaliações (**number_of_reviews**), mais comentários (**comments da planilha reviews**) e com melhores notas na avaliação geral (**review_scores_rating**)?*

Identificar características das casas com melhores avaliações nessas épocas mais movimentadas:

- *Qual é o tipo de casa (apartamento, casa, estudo...)? (**property_type**)*
- *Quantos quartos? (**bedrooms**)*

Entender mais sobre a hospedagem:

- *Qual é o tipo de locação (quarto compartilhado, quarto privativo, casa inteira...)? (**room_type**)*
- *Acomoda até quantas pessoas? (**accommodates**)*
- *Qual é a faixa de preço da diária? (**price**)*
- *Qual é a faixa de preço se for semanal? (**weekly_price**)*
- *Qual é a faixa de preço se for mensal? (**monthly_price**)*
- *Quais são as melhores notas dos preços dessas casas analisadas? (**review_scores_value**. Estou na dúvida se é esse mesmo)*
- *Tem algum diferencial na estrutura? (**amenities**)*
- *O que mais aparece na descrição dessas casas? (**description**)*

(Uma ideia para colocar no final da apresentação: como nenhum host ofereceu uma experiência para seus hóspedes, talvez fosse interessante testar algo para ser um diferencial das hospedagens. Sugestão: uma experiência poderia ser a possibilidade de descontos em passeios, visitas e restaurantes com quem o host tivesse parceria. Já que os hóspedes citaram sobre esses interesses nas avaliações sobre a localização da casa).

Sobre a localização:

- *Qual é a vibe do bairro com melhores avaliações? (**neighborhood_overview**)*
- *Qual é a nota da localização dessas casas? (**review_scores_location**)*

Com base nos dois melhores bairros, analisar:

- Contar a quantidade de casas disponíveis ao longo do tempo em cada bairro por semestre para saber se há uma tendência de crescimento muito rápido de casas sendo alugadas em temporada. Usar count os **listings_ids** totais de cada um desses dois bairros e ver quantas delas são do tipo que considerei como as melhores segundo as características identificadas na minha análise com as perguntas anteriores. Deixar um gráfico de barras empilhadas para mostrar isso.
- Fazer também um forecast dos próximos semestres, com gráfico de barras, da quantidade de casas consideradas as melhores, combinado com um gráfico de linhas com os preços médios das dessas casas por semestre.

Depois de identificar o melhor bairro:

Assim, conseguirei identificar o melhor bairro e em qual área deste bairro estão localizadas as melhores casas, segundo todos os critérios avaliados acima.

Deixar um mapa exibindo essas casas neste bairro. Usar o campo “street” para mostrar as ruas dessas casas, podendo sugerir ruas próximas e que não tenham ainda casas com essas características consideradas as melhores.

Além disso, colocar uma foto bonita desse bairro e depois mostrar o que esse bairro tem que vai satisfazer os interesses dos visitantes. Por exemplo: Mapa de metrô, localização do aeroporto, as rodovias... Pontos turísticos que esse bairro possui que vão agradar os visitantes, segundo o que a maioria avaliou como ponto positivo.

2) Limpar os dados:

- Explorar os dados e identificar o que é importante manter e o que pode ser descartado dos campos para facilitar a manipulação dos dados
- Limpar possíveis espaços
- Excluir células vazias
- Checar os tipos de dados e formatá-los quando necessário
- Padronizar dados de campos que apresentem inconsistências que possam prejudicar a análise

3) Criar tabela com principais métricas e tabelas com agrupamentos:

- Selecionar os campos definidos para a análise de acordo com as perguntas iniciais e construir uma tabela no usando SQL somente com essas principais métricas
- Criar tabelas com cálculos necessários, a partir da tabela acima, para ser base de dados agrupados para o passo 4.

4) Identificar padrões e tendências com base nas perguntas de negócio:

- Analisar o comportamento dos dados em gráficos e tabelas no Looker Studio. Isso com base na tabela feita inicialmente com as principais métricas e com os agrupamentos realizados nas pivot tables.

5) Produzir o relatório da análise

- Iniciar com o título no canto superior esquerdo, com uma foto de Boston de fundo e meu nome e profissão no canto inferior direito;
- Colocar a pergunta de negócio principal na segunda página com o mesmo tamanho do título
- No corpo da análise, colocar os gráficos e tabelas para explicar com pequenos textos os insights da análise baseada nas perguntas iniciais. Também utilizar fotos de Boston, quando conveniente.
- Finalizar com agradecimento e orientação de envio de dúvidas por e-mail.